Origin and Biologic Individuality of the Genetic Dictionary

Abstract. Deoxyribonucleic acid contains sequences complementary to homologous amino-acid transfer ribonucleic acid molecules which serve as the translating device between polyribonucleotides and proteins. This implies that the RNA molecules have their primary origin in DNA. From the amount of DNA participating, one would infer that more than 20 complementary sequences exist per genome, a conclusion consistent with a degenerate code. The fact that complex formation occurs most readily with homologous RNA suggests that, while the language remains universal, each dictionary is uniquely identifiable with its own genome.

There are excellent reasons (*1*) for identifying the soluble ribonucleic acid (s-RNA) molecules with the genetic dictionary which permits translation from the four-unit language of the polynucleotides to the twenty-unit parlance of the proteins. It is of obvious interest to determine the relation between these RNA molecules and the genome. In particular, one would like to know whether sequences exist in homologous deoxyribonucleic acid (DNA) which are complementary to those which occur in the s-RNA molecules. Information on this question would illuminate a number of central issues, including problems of origin, uniqueness of sequences, and estimations of coding degeneracy.

A test for complementarity is specific hybrid formation between DNA and RNA. This criterion was used by Hall and Spiegleman (*2*) to demonstrate that RNA synthesized in *Escherichia coli* after infection by the bacteriophage T2 is complementary to the DNA of the virus rather than to that of the host. Identification of the hybrid structures was made by equilibrium density centrifugation in swinging bucket rotors, combined with isotopic labeling. The same procedures were used (*3*) to reveal sequences in DNA complementary to homologous ribosomal RNA and to show (*4*) that DNA sequences complementary to the nucleic acid of an RNA virus do not exist. The search for complementarity with ribosomal and viral RNA made it necessary to design experiments capable of detecting complexes which include between 0.001 percent and 0.01 percent of the total DNA. The sensitivity needed was achieved by labeling the RNA to the required specific activity. Confusion with "noise" in the form of either mechanical trapping, or chance coincidence in complementarity over short regions, was avoided by making use of the resistance of specific hybrids to degradation by ribonuclease. Nonspecific complexes are completely sensitive to such degradation (*3*).

The successful detection (*3*) of hybrids of ribosomal RNA and DNA encouraged us to extend the examination of homology of DNA and RNA to the species of s-RNA molecules. The sensitivities developed for the ribosomal and viral RNA investigations (*4*) made it certain that a definitive answer was attainable for s-RNA.

Isolation, purification, denaturation of DNA, and uniform labeling of RNA with either P^{32} or tritiated uridine were carried out as described previously (*3*). Soluble RNA was obtained from a cell extract from which ribosomes were removed by centrifugation for 3 hours at 100,000*g*. The RNA was isolated by the phenol procedure (*5*) and further purified by chromatography on columns of methylated albumin (*6*). The procedures of Hall and Spiegelman (*2*) for the formation and detection of hybrid structures in cesium chloride gradients were followed.

Preliminary experiments revealed that the temperature range (40° to 55°C) suitable for complex formation with either informational (*7*) or ribosomal RNA (*3*) did not lead to hybrids between s-RNA and DNA. In one sense this was fortunate since a simple method was thus provided for detecting contamination of s-RNA with ribosomal or complementary RNA.

Hybrid formation requires incubation at temperatures and under ionic conditions that allow formation of hydrogen bonds. It is perhaps not surprising that s-RNA is resistant to hybridization since x-ray analysis (*8*) suggests that s-RNA is a hairpin structure kept together by a highly regular system of hydrogen bonding. Until this secondary structure is disrupted there is no opportunity for pairing between s-RNA molecules and complementary sequences in the DNA. One can make an educated guess of the temperatures required for complex formation by examining the melting curve of s-RNA obtained under the conditions required for hybridization.

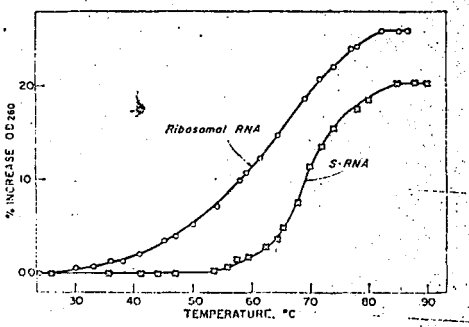

Fig. 1. Hyperchromicity of s-RNA and ribosomal RNA from *E. coli*. Solutions containing 30 μg/ml of nucleic acids in TMS (0.01*M* tris, 0.3*M* NaCl, 0.001*M* Mg^{++}, *p*H 7.3) were used. The $O.D._{260}$ readings at the ambient temperatures were made in stoppered quartz cuvettes in an Opticon spectrophotometer provided with controlled electrical heating. The temperature was increased at a rate of 0.5°C/min.

Figure 1 compares the hyperchromic response of s-RNA and ribosomal RNA to heating in the medium employed for hybridization. One would predict from these curves that complex formation between ribosomal RNA and DNA would occur between 40° and 50°C but that virtually no hybrids would be observed with s-RNA. The conclusions derived from this analysis were confirmed when the optimal conditions for hybridization of s-RNA were determined. These tests were deliberately carried out at low input ratios of s-RNA to DNA so that the conditions could be used to examine the effect of RNA concentration on extent of hybridization. Figure 2*A* shows that the amount of s-RNA hybridized per unit of DNA reached a plateau between 70° and 75°C. Figure 2*B* illustrates the kinetics of hybrid formation at 72°C and indicates that the process was nearly complete in about 100 minutes. Because of the results of these and similar experiments, mixtures undergoing hybridization were incubated at 72°C for 2 hours.

It must be emphasized that since only a minute segment of the DNA is likely to participate there will be comparatively small amounts of RNA in the apparent complex. Thus there must be independent evidence that hybrids between s-RNA and DNA are being formed. The nature of the evidence required can be listed as follows:

1) Proof that the RNA that has become a part of the complex is s-RNA by the analysis of the base composition of the hybridized material.

2) A demonstration that the RNA that has become part of the complex is much more resistant to degradation by ribonuclease than is a free RNA control.

3) Evidence that s-RNA saturates the DNA at levels indicating that only a small proportion of the DNA is complementary to the s-RNA.

Table 1. Base composition of RNA hybridized to DNA. S-RNA of *E. coli*, uniformly labeled with P^{32} to a level of 200,000 count/min per microgram, was purified on a methylated albumin column. It was incubated with homologous heat-denatured DNA in 0.01*M* tris, 0.001*M* Mg^{++}; 0.3*M* NaCl at *p*H 7.3 for 2 hours at 72°C. The input RNA/DNA was 0.16×10^{-2} to avoid nonspecific complex formation, sensitive to ribonuclease. Hybrid structures were separated as described in Fig. 2*A*. The hybrid region was pooled, and unlabeled ribosomal RNA was added as carrier. After hydrolysis in 0.3*M* NaOH for 16 hours, the nucleotide composition was determined by distribution of P^{32} in the peaks recognized on Dowex-formate columns (7). C, A, U, and G refer to cytidylate, adenylate, uridylate, and guanylate.

Moles percent				GC (%)	Purine/pyrimidine
C	A	U	G		
Input s-RNA					
27.2	20.6	18.2*	34.0	61.2	1.23
Hybridized RNA					
27.2	18.6	19.0*	35.2	62.4	1.20
Ribosomal RNA (7)					
24.3	25.0	19.7	31.0	55.3	1.27
"Informational" RNA (7)					
24.7	24.1	23.5	27.7	52.4	1.07

* Includes also counts attributable to pseudouridylate.

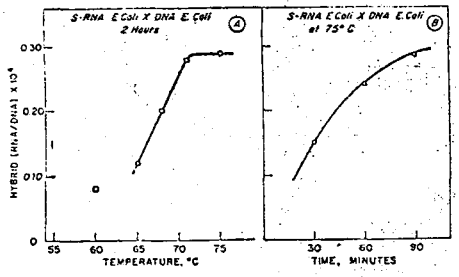

Fig. 2. Hybridization s-RNA to DNA. (*A*) Influence of temperature. The incubation mixture in TMS consisted in every case of 40 μg of heat-denatured *E. coli* DNA and 0.01 μg of H^3-labeled s-RNA *E. coli* (10^5 count/min per microgram). The total volume was 0.6 to 0.7 ml. The mixtures were incubated for 2 hours at the specified temperatures. Enough solid or solution of CsCl was then added to the mixtures to yield a density of 1.72, and a final volume of 3.0 ml. The resulting samples were centrifuged for 3 days at 25°C and 33,000 rev/min in the SW39 rotor of a Spinco model L ultracentrifuge. After centrifugation 0.2 ml fractions were collected from the bottom of the tubes, diluted to 1.2 ml and analyzed for $O.D._{260}$ and radioactivity of the acid-precipitable fraction (*3*). The amount of RNA hybridized is determined from the counts of the acid-precipitable fraction in the DNA-density region of the gradient. The input ratio was adjusted so that nonspecific, ribonuclease-sensitive counts were not significant. (*B*) Influence of time. The experimental conditions described in part *A* were used, except that here the temperature was kept constant at 75°C and the time of incubation was varied.

4) Specific complex formation with homologous and but not heterologous RNA.

To satisfy the first point, purified s-RNA, uniformly labeled with P^{32}, was hybridized to DNA. After separation in a CsCl gradient, the hybridized material was removed, hydrolyzed by alkali, and analyzed for base composition (*7*). Table 1 shows the base composition of the hybridized material and the input s-RNA. For comparison, the base compositions of the other RNA components of *E. coli* are also listed. The hybridized RNA is virtually identical in base composition to the s-RNA and is easily distinguishable from either the ribosomal or the "informational RNA" which is synthesized in a "step-down transition" (*7*). It should be noted that pseudouridylate, a base unique to s-RNA, was present in its characteristic position on the Dowex-formate column.

The second property, the relative resistance to ribonuclease of s-RNA (H^3-labeled) hybridized to DNA is illustrated in Fig. 3. Here P^{32}-labeled s-RNA was added to the reaction mixture as an internal control to monitor the enzyme activity. The initial loss in acid-precipitable material in the tritium counts is due to adventitious s-RNA contaminating the DNA-density region. As might be expected the proportion varies with the input ratio of s-RNA to DNA; it is very small at low input ratios (see Fig. 4).

The next question refers to the level at which saturation of the DNA occurs. A fixed amount of DNA was incubated with various concentrations of tritiated s-RNA and the resulting complexes were separated in CsCl gradients. The total and the ribonuclease-resistant material in the DNA density regions were then determined. The results are summarized in Fig. 4. The picture is very similar to that observed in the saturation experiments with ribosomal RNA (*3*). Before treatment with ribonuclease no sharp plateau is observed owing to the formation of irrelevant complexes when the ratio of RNA to DNA is high. However, counts of the ribonuclease-resistant material show abrupt evidence of saturation. The plateau suggests that 0.023 percent of the DNA consists of sequences complementary to s-RNA.

We now turn to the last criterion, specificity of complex formation. Convincing tests of specificity, which contain internal controls on the hybridizing process, have been designed. In the

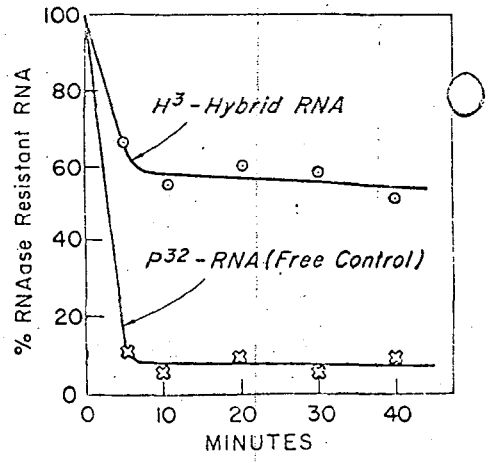

Fig. 3. Resistance of "hybridized" RNA to ribonuclease action. Hybridization was carried out with H^3-labeled s-RNA and heat-denatured DNA of *E. coli* in TMS at 72°C for 2 hours. The input RNA/DNA was 2.6×10^{-2} to permit exhibition of some nonspecific complex in the DNA-density region. After density centrifugation, as described in Fig. 2, the fractions containing hybrid were pooled and dialyzed against a solution containing 0.01*M* Tris and 0.1*M* NaCl at *p*H 7.3. Free P^{32}-labeled s-RNA was included as an internal control. Ribonuclease (10 μg/ml) was added and the incubation was continued at 25°C. At the indicated intervals, samples were removed and the H^3- and P^{32}-counts of the acid-precipitable residue were determined in a Packard scintillation spectrophotometer (*3*).

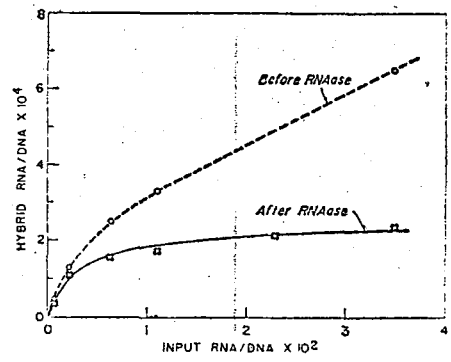

Fig. 4. Saturation of s-RNA *E. coli* by DNA *E. coli*. Mixtures containing 45 to 50 μg of heat-denatured DNA plus various amounts of H^3-labeled s-RNA (100,000 count/min per microgram) in 0.85 ml TMS were incubated at 72°C for 2 hours. The samples were then treated as described in Fig. 2*A*. After separation in the density gradient, the DNA-density fractions were collected and divided into two portions. In one, the radioactivity of the total acid-insoluble material was determined. The second portions were further diluted so that the concentration of CsCl would be less than 0.2*M* and then 10 μg/ml of ribonuclease was added. After incubation at 25°C for 10 minutes the radioactivity of the acid-precipitable residue was determined. In the plateau region, the whole hybrid fraction was pooled and the ribonuclease-resistant residue was examined by the kinetic procedure of Fig. 3.

first, s-RNA preparations of two different biological origins were incubated with DNA homologous to one of them. For ease of identification, each s-RNA is labeled with a different radioactive isotope. If specificity is complete, the homologous s-RNA will hybridize and the heterologous RNA will be excluded. In the other type of test, two DNA preparations, distinguishable by their positions in the density gradient, were incubated in the presence of a labeled s-RNA homologous to one of them. If specificity does exist, the isotope should be found associated only with the peak corresponding to the homologous DNA.

The results of both kinds of specificity tests are summarized in Fig. 5. In Fig. 5*A* the incubation mixture contained denatured DNA from *E. coli*, P^{32}-labeled s-RNA of *E. coli* and H^3-labeled s-RNA of *Bacillus megaterium*. We note complete exclusion of the tritiated heterologous s-RNA from the hybrid region and excellent hybrid formation with the homologous s-RNA marked with P^{32}.

The alternative specificity test was carried out with DNA of *Pseudomonas aeruginosa* and DNA of *Bacillus megaterium* since they have very different densities and separate well in CsCl gradients. A mixture of heat-denatured preparations of DNA from these two sources were incubated with tritium-labeled s-RNA of *B. megaterium*. The resulting profiles obtained in the CsCl gradients are shown in Fig 5*B*. There are virtually no counts associated with DNA from *Pseudomonas aeruginosa* whereas an excellent hybrid structure is seen in the density region of the homologous DNA from *B. megaterium*.

The experiments described offer convincing evidence that specific hybrids between s-RNA and DNA can be formed under the proper conditions. The material which has formed a complex is indeed s-RNA as shown by a direct analysis of the hybridized material. The usual resistance of hybrids of RNA and DNA to ribonuclease has been demonstrated. Further, the DNA is saturated at RNA levels that indicate only a small proportion of the DNA is complementary to the s-RNA molecules. Finally, hybrid formation is specific, occurring readily only between homologous s-RNA and DNA.

A number of interesting implications follow from the fact that DNA contains sequences complementary to s-RNA. The existence of complementarity makes it likely that s-RNA originates on a DNA template. Invoking a DNA-independent pathway for the origin of s-RNA (*8, 9*) seems now unnecessary. This conclusion is in agreement with the recent demonstration that actinomycin *D*, which inhibits (*10*) the DNA-dependent RNA polymerase, also prevents (*11*) all RNA synthesis in both bacterial and animal cells. The fact that this same agent does not inhibit production of RNA virus (*11*) is argument against the operation in uninfected cells of a mechanism which synthesizes RNA on an RNA-template.

The presence in s-RNA of methylated bases and pseudo-uridine might perhaps pose a problem for a DNA-dependent pathway. However, methylation of purines and pyrimidines occurs after the formation of the polynucleotide (*12*). We would then predict the existence of an ezyme which can convert uridine to pseudo-uridine in the intact polynucleotide.

The fact (from the hybridization curve) that s-RNA saturates the DNA at about 0.02 percent is consistent with what would be expected if the genetic code were degenerate. Despite the complicated operations required to obtain this number it is remarkably reproducible. In independent experiments values of 0.019 percent and 0.023 percent were obtained in this laboratory. In addition, Goodman and Rich (*13*) in a very similar experiment found a value of 0.024 percent. Therefore, the value can be given some credence. If anything, it is probably an underestimation since conditions for hybridization may not be optimal and the treatment with ribonuclease may slowly remove some s-RNA that had entered into the complex.

The expected saturation level can be estimated from the molecular weight equivalent of the genome of *Escherichia coli* (4×10^9) and the number of different kinds of s-RNA molecules, each of which has a molecular weight of 2.5×10^4. If the genetic code is not degenerate, each amino acid is coded by only one triplet which implies that there are only 20 different s-RNA molecules. If the code is degenerate, more than 20 will be needed in the dictionary. The plateau predicted by the non-degenerate case is 0.01 percent. The fact that it is at least twice as high suggests that some amino acids are identified with more than one s-RNA molecule. This possibility is consistent with the accumulating evidence for degeneracy which has emerged from triplet identifications (*14, 15*) and agrees with the multiplicity of types of s-RNA for individual amino acids (*16, 17*). That this multiplicity is the physical basis for degeneracy has been demonstrated by Weisblum *et al.* (*18*).

The data available (*19, 20*) suggest that the genetic dictionary is universal, or nearly so. However, the coding triplets probably occupy only a small proportion of the s-RNA strands. Although the function of the non-coding stretches of approximately 70 nucleotides is as yet unknown. They provide the opportunity for biological individuality by sequence variation without disturbing the functioning of the universal language. The specificity of complex formation in the present experiments shows that this opportunity was not neglected in the course of biologic evolution. Thus, although the s-RNA of *E. coli* can translate the genetic message of a rabbit into hemoglobin (*19*), the s-RNA can be uniquely identified with the genome of its origin.

Ribosomal RNA appears to have the

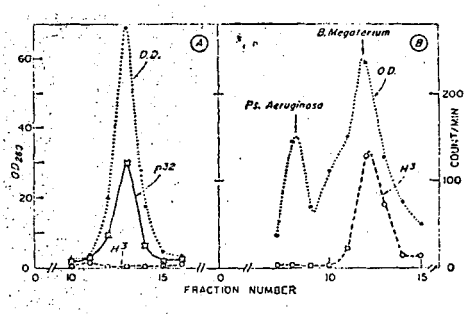

Fig. 5. Specificity of hybridization. (*A*) Two different types of s-RNA plus one type of DNA; 55 μg of heat-denatured DNA from *E. coli* were incubated in TMS (2 hr, 72°) with a mixture of 0.06 μg of P^{32}-labeled s-RNA (50,000 count/min per microgram) from *E. coli* and .06 μg H^3-labeled s-RNA (100,000 counts/min per μg) from *B. megaterium*. After equilibrium density gradient centrifugation (Fig. 2) 0.2 ml fractions were collected and diluted to 1.2 ml. The $O.D._{260}$ and the radioactivity of the ribonuclease-resistant acid-precipitable fraction were determined by counting in a Packard liquid scintillation spectrometer which permits simultaneous counting of P^{32} and H^3. Only the hybrid region of the density gradient is shown. The optical-density profile identifies the DNA and the radioactivity identifies the hybrid structure. (*B*) Two different DNA's plus one s-RNA. Heat-denatured DNA from *B. megaterium* and *Ps. aeruginosa*, 30 μg each, were incubated in TMS with 0.06 μg of H^3-labeled s-RNA from *B. megaterium* at 72°C for 2 hours (total volume was 0.7 ml). The same procedure as in Fig. 2*A* was followed. The hybrid portion of the curve after treatment with ribonuclease is shown. The peaks in the optical-density profile identify the positions of the two DNA preparations in the reaction mixture.

same combination of genetic uniqueness and use unrestricted by specificity requirements. Ribosomes are comparatively indifferent (*14–16*) to the origin of the genetic messages to which they respond. However, their sequences are unique, since they hybridize readily only to homologous DNA (*3*).

One other feature is shared by these two molecular species. Although their sequences vary, the over-all base composition is remarkably similar in a variety of organisms. Thus, ribosomal RNA from *Pseudomonas aeruginosa* with a DNA containing 64 percent guanine-cytosine is indistinguishable from that of *Bacillus megaterium*, the DNA of which contains 44 percent guanine-cytosine. We are faced with the paradox that two sets of cistrons, those for s-RNA and ribosomal, have resisted the drift toward different average base compositions (*21*).

DARIO GIACOMONI
S. SPIEGELMAN
Department of Microbiology, University of Illinois, Urbana

References and Notes

1. M. W. Nirenberg, J. H. Matthaei and O. W. Jones, *Proc. Natl. Acad. Sci. U.S.* 48, 104 (1962); F. Chapeville, F. Lipmann, G. von Ehrenstein, B. Weisblum, W. J. Ray, Jr., S. Benzer, *ibid.* 48, 1086 (1962).
2. B. D. Hall and S. Spiegelman, *ibid.* 47 137 (1961).
3. S. A. Yankofsky and S. Spiegelman, *ibid.*, 48, 1069 (1962); *ibid.*, 48, 1466 (1962).
4. R. H. Doi and S. Spiegelman, *Science*, 138, 1270 (1962).
5. A. Gierer and G. Schramm, *Nature*, 177, 702 (1956).
6. J. D. Mandell and A. A. Hershey, *Anal. Biochem.* 1, 66 (1960).
7. M. Hayashi and S. Spiegelman, *Proc. Natl. Acad. Sci. U.S.* 47, 1564 (1961).
8. J. Spencer, W. Fuller, M. H. F. Wilkins, G. L. Brown, *Nature* 194, 1014 (1962).
9. D. S. McCully and G. L. Cantoni, *J. Mol. Biol.* 5, 80 (1962).
10. J. Hurwitz, J. J. Furth, M. Malamy, M. Alexander, *Proc. Natl. Acad. Sci. U.S.* 48, 1222 (1962).
11. E. Reich, R. M. Franklin, A. J. Shatkin, E. L. Tatum, *ibid.* 48, 1238 (1962).
12. E. Fleissner and E. Borek, *ibid.* 48, 1199 (1962).
13. H. M. Goodman and A. Rich, personal communication.
14. R. G. Martin, J. H. Matthaei, O. W. Jones, M. W. Nirenberg, *Biochem. Biophys. Res. Commun.* 6, 410 (1962).
15. J. F. Speyer, P. Lengyel, C. Basilio, S. Ochoa, *Proc. Natl. Acad. Sci. U.S.* 48, 441 (1962).
16. P. Berg and V. Lagerkvist, *Acides Ribonucleiques et Polyphosphates*, (C.N.R.S., Paris 1962).
17. B. P. Doctor, J. Apper, R. W. Holley, *J. Biol. Chem.* 236, 117 (1961); N. Sueoka and T. Yamane, *Proc. Natl. Acad. Sci. U.S.* 48, 1454 (1962).
18. B. Weisblum, S. Benzer, R. H. Holley, *ibid.* 48, 1449 (1962).
19. G. von Ehrenstein and F. Lipman, *ibid.* 47, 941 (1961).
20. A. Tsugita, H. Fraenkel-Conrat, M. W. Nirenberg, J. H. Matthaei, *ibid.* 48, 846 (1962); B. Weinstein and A. N. Schechter, *ibid.* 48, 1686 (1962); E. S. Maxwell, *ibid.* 48, 1639 (1962).
21. Aided by grants from the U.S. Public Health Service, National Science Foundation, and the Office of Naval Research. One of us (D.G.) is an Upjohn predoctoral fellow in molecular genetics.

16 November 1962